Economics is not what you think: A defense of the economic approach to taxation
Marc Fleurbaey

To cite this version:
Marc Fleurbaey. Economics is not what you think: A defense of the economic approach to taxation. 2012. halshs-00698575

HAL Id: halshs-00698575
https://halshs.archives-ouvertes.fr/halshs-00698575
Preprint submitted on 16 May 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Economics is not what you think: A defense of the economic approach to taxation

Marc Fleurbaey

N°10 | may 2012

In *The Myth of Ownership - Taxes and Justice*, Liam Murphy and Thomas Nagel (2002) launch an attack against a straw man, the economist who believes that taxation should minimally interfere with property rights and should seek to preserve the market distribution of wealth and income. Instead they propose an approach that focuses on the consequences of any form of public intervention for the distribution of welfare, without any particular ethical concern for the values emerging from the market. In fact, such an approach has been long developed by Mirrlees (1971), whose approach has been dominating the economics of taxation for the last forty years. But more recently the fairness approach to taxation goes beyond welfare consequentialism and attributes some value to market allocations, in line with the theories of justice proposed by Rawls and Dworkin.
Economics is not what you think: A defense of the economic approach to taxation

Marc Fleurbaey

May 2012

About the author
Marc Fleurbaey is Laurance S. Rockefeller Professor at the University of Princeton (USA). His research in economics and philosophy has been devoted to normative economics, distributive justice, and the evaluation of public policy. The main areas of application are the measurement of social welfare, climate policies, health prioritizing, and optimal taxation.

He has been professor of economics and CNRS researcher in France. He is the coordinating editor of Social Choice and Welfare and a former editor of Economics and Philosophy. He is a Coordinating Lead Author for the Fifth Assessment Report (2014) of the Intergovernmental Panel on Climate Change. He has been a member of the Stiglitz-Sen-Fitoussi Commission on the Measurement of Economic Performance and Social Progress. He has published Fairness, Responsibility, and Welfare (Oxford 2008) and A Theory of Fairness and Social Welfare (Cambridge 2011, with F. Maniquet), as well as many articles in welfare economics, public economics, and political philosophy.

The text
This text was written by Marc Fleurbay, Princeton University, who holds the Chaire entitled « Welfare Economics and Social Justice » at the Collège d’études mondiales.


Reference to this document
Marc Fleurbaey, Economics is not what you think: A defense of the economic approach to taxation, FMSH-WP-2012-10, may 2012.
Abstract
In *The Myth of Ownership – Taxes and Justice*, Liam Murphy and Thomas Nagel (2002) launch an attack against a straw man, the economist who believes that taxation should minimally interfere with property rights and should seek to preserve the market distribution of wealth and income. Instead they propose an approach that focuses on the consequences of any form of public intervention for the distribution of welfare, without any particular ethical concern for the values emerging from the market. In fact, such an approach has been long developed by Mirrlees (1971), whose approach has been dominating the economics of taxation for the last forty years. But more recently the fairness approach to taxation goes beyond welfare consequentialism and attributes some value to market allocations, in line with the theories of justice proposed by Rawls and Dworkin.

Keywords
taxation, welfare, property rights, fairness, Mirrlees

Résumé

Mots-clés
fiscalité, bien-être, droits de propriété, équité, Mirrlees
In their celebrated *The Myth of Ownership – Taxes and Justice*, Liam Murphy and Thomas Nagel (2002) launch an attack against a straw man, the economist who believes that taxation should minimally interfere with property rights and should seek to preserve the market distribution of wealth and income. Instead they propose an approach that focuses on the consequences of any form of public intervention for the distribution of welfare, without any particular ethical concern for the values emerging from the market.

The straw man does exist in some corners of the public debate, although such corners, in the United States, are now moving farther to the libertarian extremes of condemning any form of taxation as evil. And it is also true that the old public finance literature on taxation had little to say apart from vague principles of horizontal and vertical equity. In particular, as Murphy and Nagel rightly argue, the vertical equity principle has remained empty for a long time because beyond the idea that the better endowed should pay more, it did not say much, and in particular it did not say how much more. The equal sacrifice principle was sometimes invoked, but, as they rightly say again, there is little reason to seek to preserve the distribution of utility that comes out of the unfettered market, so that this principle appears groundless.

The straw man no longer exists in academic economics, however. After James Mirrlees (1971) introduced the incentive compatibility approach to the taxation problem, combined with a utilitarian social welfare function, what economists do about the taxation problem is exactly what Murphy and Nagel suggest should be done. Economists invoke an objective defined in terms of a final distribution of utilities, and they try to determine what tax system would induce the best distribution after individual agents react to the incentives laid out by the tax formula. Murphy and Nagel simply ignore the Mirrlees approach until, around page 135, they discuss the issue of the optimal progressivity of the tax rate. Less than four pages of the book are devoted to the Mirrlees approach, in spite of the fact that it has been dominating the economic literature on taxation for the last forty years.

The paradox is that the Mirrlees approach, arguably, went too far in the direction later advocated by Murphy and Nagel. By relying on a utilitarian social welfare function, or some generalized form of it, it embraces a narrow form of consequentialism that focuses on utility and completely disregards the way in which utility is generated. At the same time as Mirrlees was publishing his seminal contribution, John Rawls published another seminal work that sparked interest for the distribution of resources and for a moral division of labor between individuals and society. It is not society's job, according to Rawls, to maximize a certain distribution of utility, because individuals have to assume some responsibility for their own utility.

After Rawls, several philosophers and most notably Dworkin, Arneson, and Cohen have developed this idea of a division of labor in which individual responsibility plays a key role, and that came to be called the luck egalitarian view. Economists have not ignored this new line of thought, and in fact they have recreated it more or less independently. As it turns out, in this decidedly special year of 1971, Serge Kolm published the working paper version of a (seminal) monograph on equity that generated a literature on fairness in the distribution of resources. In this approach, fairness principles govern the distribution of resources and, while individual preferences are taken into account so that the allocation may be efficient, no attention is paid to utilities, which remain a private matter. This literature, until recently, said very little about the taxation problem because it focused on the “first-best” context, i.e., the case of a perfectly planned economy with no incentive problem. But applications to taxation have now been made, which deliver a somewhat different outlook from the Mirrlees-Murphy-Nagel approach.

The purpose of this chapter is to introduce readers to the economic approach to taxation, in the Mirrlees tradition and in the more recent fairness approach. It is hoped that, from this non-technical introduction to the economic approach, economists will get a better press than after Murphy and Nagel's charge against a straw man. The conclusion of the chapter then comes back to the interesting question of whether market values have a special moral significance.

1. The link between philosophy and economics is complex here. Kolm (1972) quoted Rawls but gave a welfarist interpretation of the difference principle. More to the point, Varian (1975) explained how Rawls' idea of equality of resources was embodied in some of the economic concepts. But the general motivation of the economic study of fair allocation in the 1980-90s did not appear to come from the philosophical literature.
An important restriction will be adopted throughout the chapter. It will be assumed that the only source of income for individuals is labor earnings, and only income taxation will be examined. There is obviously a literature on other forms of taxation (capital, commodity, inheritance), but it will be simpler to focus here on the taxation of labor earnings.

**Mirrlees and optimal taxation**

The incentive problem

There are two important features in the approach initiated by Mirrlees. The first is a rigorous analysis of the incentive problem. It was long known that taxing earnings may have a discouraging effect on labor supply, although the effect is on the whole uncertain because the fact that a tax impoverishes people may induce them to work more to keep their consumption level. It is a standard balancing act in consumer-worker theory between a substitution effect (the tax makes leisure less expensive in foregone consumption, therefore they work less) and an income effect (the tax makes one poorer, and people may want to share this scarcity between a lower consumption and a lower leisure, therefore they work more).

What is new in the Mirrlees approach is the understanding that when only earnings and taxes are observed, and not the effort and the labor hours that generate earnings, it is impossible to distinguish people who earn the same earnings with different quantities of work or effort. The important fact is then that the more talented an individual, the easier it is for her to obtain any given level of earnings. Therefore, those who have better hidden talents (i.e., who have access to more rewarding jobs) are necessarily better off than the others, for given preferences. Incentives act as a constraint on redistribution. It is simply impossible to equalize utilities, unless one neutralizes the impact of effort on net income by imposing a 100% marginal tax rate—a recipe for economic collapse.

This problem is now described in economic theory as the "self-selection" constraint. Public policy amounts to proposing a menu of earnings and taxes to individuals, and they choose what they prefer in this menu. It is impossible to make some of the options accessible only to low-talented individuals, because the more talented individuals can always pretend that they have low talent and they can therefore mimic the low-talented by choosing the same options.

Murphy and Nagel criticize the Mirrlees approach for emphasizing the incentive problem. They argue that labor supply is rather insensitive to taxes for most people. They also consider the view that declarations of earnings, not just earnings levels, may be sensitive to taxes, but then note that it is a problem more for the design of a good tax monitoring administration than a real incentive problem. The incentive problem is, however, more serious than they acknowledge. First, labor supply is sensitive to taxes for the second breadwinner more than the first one in many households, for which the choice is not just to work a few hours less but also to move from full time to part time, or from part time to home-making. Second, the issue is not so much about labor hours but also about effort at work. When marginal taxes are high, people may feel less eager to exert a lot of energy to improve their career track. Third, people may also be affected in their choice of jobs. High taxes may induce people to choose jobs with greater intrinsic rewards and less monetary payoffs, as in the choice between academia and the finance sector for talented people.

In any case, the real problem with the incentive constraint is not so much that people would earn much less if taxes were somewhat higher, but that it is hard to achieve substantial equality because private information about personal talents give an edge to the better endowed.

The optimization

The second important feature in Mirrlees’ approach is the technical feat of solving a double optimization program. The difficulty is indeed that one should choose a tax formula that maximizes the value of a social welfare function, under the constraint that every individual in the population will solve his or her own optimization problem when choosing his or her own level of earnings. At the age of computers, this achievement may appear less important because one can always incorporate this double problem into a program and obtain a numerical solution. But finding a theoretical solution provides a useful way to understand how the optimal tax formula depends on the data of the problem, namely, the distribution of skills in the population, the preferences, the degree of priority for the worse-off
in the social welfare function. After Mirrlees, the theoretical solution has found a simpler and elegant formulation in the work of Atkinson and Stiglitz (1980), Diamond (1998) and Saez (2001).

An issue that has attracted nervous interest is the theoretical result that the individuals with the greatest skill (i.e., those who have access to the greatest wage rate in the labor market) should be submitted to a zero marginal tax rate, i.e., should pay no tax on the last dollar earned. This has appeared in contradiction with the tradition of increasing marginal tax rates for high incomes. This result, however, is not particularly shocking. Even with declining marginal tax rates over the whole range of earnings, it is possible to have increasing average tax rates. And the redistributive efficacy of the tax system depends on increasing average rates, not on the marginal rates. To illustrate, consider the simple case of a constant marginal tax rate. When the tax formula involves a negative tax for low incomes, the average tax rate is initially negative and then slowly increases toward the level of the marginal rate. Therefore, when the marginal tax rate decreases it is also possible to have constantly increasing average rates. It is true, though, that when the marginal tax rate falls to zero, the average tax rate will fall at some point for very high incomes, but the range over which it falls may be quite small.

Indeed, simulations have been made and showed that the range over which the marginal (not just the average) tax rate falls may be small and affect only the highest incomes (Tuomala 1990). Diamond (1998) and Saez (2001) noted that if the upper tail of the distribution of skills is fat enough, marginal tax rates never fall (this, however, requires that there be infinite skills).

Qualities of the approach, and two limitations
In summary, the Mirrlees approach to optimal taxation has the appealing feature that it evaluates the tax policy by its consequences on the distribution of well-being, not by any narrow principle of tax fairness. This is important not just because it avoids the flaws and indeterminacy of the vertical equity approaches criticized by Murphy and Nagel, but also because it makes the analysis of tax policy congruent with the analysis of any other public policy. If education policy, health policy, and so on, are evaluated with the same social objective as the tax policy, one obtains a consistent setting in which all instruments are aimed at the same direction.

Another nice feature of the approach is that the social welfare function is increasing in each individual’s utility, so that it fully embodies the Pareto principle. At the same time, it may incorporate aversion to inequality, or a degree of priority to the worse off. This did not appear in Mirrlees’ original formulation, which adopted the sum-utilitarian objective, but soon other authors have explored other objectives, including the maximin that gives absolute priority to the worse off (Atkinson 1975). This combination of the Pareto principle and a priority for the worse off means that one is not trapped in the dilemma of having to choose between efficiency and equity. Both concerns will be optimally balanced in the selected policy. The policy will be efficient given the feasibility constraints, in particular the incentive constraints, but it will also take care of the worse off and, again, produce the best distribution that is possible given the feasibility constraints.

Of course, economists may still want to compute the deadweight loss due to taxation, as compared to the laisser-faire allocation. But, while the laisser-faire is in the set of feasible allocations (“no tax” is a particular tax policy), the objective function does not give any particular role to the market. The distribution of well-being is the only argument of the social welfare function, and this same function could be used in a planned economy just as well as in a market economy. The institutions are considered mere instruments in the pursuit of the welfarist objective.

There is, however, an important limitation to the Mirrlees tradition. It generally assumes that the individuals have the same preferences, so that the same utility function can be used to represent their interests in the social welfare function. Attempts to allow for heterogeneous preferences in the population have generally involved arbitrary weights for each type of utility function featuring in the social welfare function, and therefore no specific result about the desirable direction of redistribution.

Introducing multiple preferences in the optimization problem also makes the incentive problem intractable. While there is no difficulty to extend the definition of the self-selection constraint to this case, and no difficulty to see that incentive compatibility is still satisfied when the same tax
formula is offered to all and the tax depends only on their earnings, the problem is that the analysis of incentive compatible allocations becomes complex. In a nutshell, it is then hard to guess who will want to mimic whom. Let us explain why.

To do this, let us first look at the case of uniform preferences, which is much simpler. The better skilled are then always less averse to earning more (i.e., they need less extra consumption to accept earning an additional dollar), and if the social objective is redistributive, we know that the constraint is that the better skilled may want to mimic the less skilled. In order to avoid that, the less skilled must be induced to work less than would be efficient at their wage rate, a move that appears less attractive to the better skilled who are less averse to earnings and therefore less interested in leisure. Therefore the distortion induced by the tax is well understood.

In contrast, with diverse preferences, making the less skilled work less will make their situation more attractive to the talented who are, by their different preferences, strongly averse to work. Incentive constraints may then operate in all directions of skills. The structure of the set of incentive compatible allocations is then hard to analyze, and no theoretical solution has yet been found to the double optimization problem in this context. In order to avoid this difficulty, authors routinely assume that the distribution of skills and preferences is such that the heterogeneity of individuals can still be described as depending on a single parameter, which is not a realistic solution.

But the problem of diverse preferences is not just about feasibility and incentives. It is also normative, because the social objective must then involve interpersonal comparisons of individuals with diverse preferences. In the welfarist tradition of welfare economics and social choice theory, there is no principle on which such comparisons can be grounded. This tradition always assumes that the relevant utility functions are provided by some external authority. Economists who want to say something about the optimal tax therefore can only consider the various possible results that can come for the various possible weights attached to different utility functions.

Saez (2001, 2002) has proposed an ingenious way to address the difficulty. It consists in a considering a social welfare function that directly evaluates people’s situation in terms of net income. The degree of priority assigned to people earning a certain level can then be viewed as the average weight that a standard social welfare function would assign to the various categories of people (with different skills and preferences) earning that level. The problem is that this works well only if one can assume that the people earning a certain level of income would react similarly to a change in the tax formula. If they are heterogeneous in skills and preferences, however, this assumption cannot be made, unless heterogeneity boils down to a single parameter, as in a related approach described two paragraphs earlier.

In conclusion, one can say that the problem of multiple preferences is an important limitation of the Mirrlees tradition. It calls for a deeper normative analysis of the social objective. As it turns out, the diversity of preferences is precisely the main focus of the fair allocation approach introduced in the next section.

The fair allocation approach, when it is compared to the Mirrlees approach, also reveals that a complete neglect of the market in the objective that is maximized may be excessive. Indeed, in a perspective of a division of labor between social institutions and individual responsibility, fairness principles may bear on the distribution of transfers between individuals with certain characteristics, not just on the distribution of final well-being. The neglect of transfers in the objective function is, perhaps, another important limitation of the Mirrlees approach. This is more debatable and will be discussed in more detail in the following sections.

**Fair optimal taxation**

**Fair allocation theory**

The reader who is familiar with Rawls’ and Dworkin’s theories of justice (Rawls 1971, Dworkin 2000) knows that one can conceive the just institutions as focusing on the distribution of resources rather than well-being. Once the resources are suitably distributed, individuals may use them as they wish, and derive whatever welfare they want or can from them. The transformation of resources into well-being is considered a private matter, not an issue of social justice. Recall that personal talents can be treated as internal resources that must be counterbalanced with transfers in external resources (as in Dworkin’s approach), so that inequalities in capacities...
to transform external resources into well-being are not always neglected in this approach.

This approach to social justice has raised many objections, of course, but the emphasis in this chapter will simply be on the similarity between this approach and the economic theory of fair allocation. The economic theory of fairness has initially been developed around the concept of no-envy. The idea is to produce a situation in which no one would prefer to have the resources (including, possibly, the internal resources) of anyone else. One of the important observations made by economists has been that, when there are no differences in internal resources, i.e., when all relevant resources are transferable, then a competitive market in which everyone has the same wealth not only guarantees no-envy (because everyone has the same budget constraint, therefore the same opportunities for resources, and could have chosen whatever any other individual has) but is the only way to achieve no-envy when there is sufficient diversity of preferences.

An egalitarian market allocation, however, is not the only way of conceiving equality of resources. It has some important drawbacks, most notably with respect to solidarity with respect to changes in general circumstances. Suppose that the available resources increase, for instance. Then the market moves to another allocation, which new prices. If the relative prices of some goods go up, then the individuals who like these goods more than other individuals may end up being worse off than in the former equilibrium which had less total resources. This may occur in spite of keeping the full equality of budgets between all individuals. It appears undesirable that a change in general circumstances that is potentially good for all may hurt some individuals.

Solidarity is easier to preserve when one resorts to another approach to resource equality, named egalitarian-equivalence. This alternative approach consists in seeking to achieve a situation that gives individuals the same utility level as an allocation in which they would consume the same bundle of goods, or, in a more general form of the approach, in which they would choose from the same opportunity set. This may sound more welfarist than the egalitarian market, but it is not really. In fact utilities play no role here, only ordinal preferences matter. The goal is to have every individual being made indifferent between her current bundle and some bundle (or opportunity set) that serves as the reference for all.

It is easy to see how this guarantees solidarity. When more resources become available, the reference bundle (or opportunity set) can be improved for all, which implies that everyone's satisfaction will necessarily raise. But no-envy is not generally guaranteed. It may even happen that what Jack deems equivalent to the reference bundle contains more of every good than what Jill deems equivalent to the reference bundle. Therefore, resource equality in the sense of egalitarian-equivalence is compatible with some consuming more of every good than others. This is not particularly shocking because this is still equivalent to everyone consuming the same bundle. But what it suggests, and which is indeed true as a general theorem, is that it is impossible to combine no-envy and solidarity into a single approach that also satisfies the Pareto principle. (It is always possible to satisfy no-envy and solidarity by giving the same bundle to everyone, but this ignores the diversity of preferences and is therefore inefficient.)

Let us now see how fair allocation theory applies to the Mirrlees problem of redistributing income between individuals who have different productivities (i.e., different earnings per hour worked). Unlike Mirrlees, let us allow for heterogeneous preferences as well. In order to keep this chapter short, we will focus on egalitarian-equivalence and ignore the no-envy perspective.

An individual who works full time and earns $4,000 a month might feel equally well-off if he didn't have to work and got a lump-sum support of $2,500. This equivalent amount of leisurely income can be used as a measure of advantage, and one can seek to obtain an efficient allocation in which equality is achieved with respect to such equivalent amounts. For instance, the allocation could have low-paid workaholics who work more than full time and earn $3,000, as well as work averse but talented individuals who earn $4,000 by working part time, and many intermediate situations, but all would be indifferent between their situation and getting a certain quantity $X without having to work.

A nice feature of such an allocation is that individuals who have the same preferences is that they mutually consider equally

2. The fair allocation approach described in this section is developed in great detail in Fleurbaey (2008, chap. 4-5).
good, even when their earnings potential differs widely. Therefore this neutralizes the differences in earning abilities, which may be deemed unfair if they come from individual circumstances for which individuals should not be held responsible.

Another nice feature of this particular allocation is that it does not penalize the individuals who are work averse. On the contrary, they typically end up at situations that are advantageous in terms of income and leisure. This is because, in order to make them as well off as at the same $X (with no labor) as the others, they must either earn more than the others or work less. It is nice not to penalize work averse individuals when one suspects that their preferences are influenced by family responsibility and time constraints.

This is only an example and many variants can be imagined, including variants that hold individuals partly responsible for their earning ability. One such variant deserves to be introduced because it is relevant to the discussion about the moral value of the market. It consists of taking as reference not a situation in which the individual does not work, but a situation in which he obtains a lump-sum transfer $X and can work at the minimum wage. For instance, our individual who earns $4,000 working full time might feel equally satisfied with working part time, earning $1,000 at the minimum wage, and getting $2,130 as a lump-sum transfer every month. The goal is then to obtain an allocation in which this equivalent lump-sum transfer (associated with the possibility to work at the minimum wage) is the same for all individuals.

The reason why the minimum wage is taken as the reference is, again, to avoid punishing the work averse individuals. The greater the reference wage, the more one obtains an allocation in which work averse individuals obtain low income. The previous example that involved zero work is, obviously, even more favorable to work averse individuals, and could be described as based on a null reference wage (being offered the opportunity to work at a null wage rate, anyone with the slightest aversion to work would choose not to work).

From transcendental to comparative
These are just two examples of the egalitarian-equivalence approach applied to the income redistribution problem. The advantage, compared to the Mirrlees approach, is that a diversity of preferences is possible. But a very important limitation of the examples provided in the previous subsection is that they have neglected the incentive constraints. Equalizing the equivalent lump-sum transfers, and giving more income and more leisure to some individuals just because they have greater aversion to work, is impossible if preferences are private information. Everyone would pretend to be work averse if this were a way to obtain a better tax treatment.

In order to cope with incentive constraints, the most convenient method is to reformulate the objective into a ranking of all possible allocations. In Sen’s (2009) terminology, the theory of fair allocation described so far is “transcendental”, describing an ideal state of society that is totally efficient and totally fair at the same time. But what one needs is a “comparative” approach that evaluates all possible allocations and enables the policy-maker to choose the best in the set of feasible allocations that is delineated in particular by incentive constraints.

The transformation of the theory of fair allocation into a theory of fair social orderings has been initiated in the last decade and is still in progress3. An advantage of the egalitarian-equivalence approach, in contrast with the no-envy approach, is that it lends itself naturally to such a transformation. Indeed, individual situations are evaluated in terms of an equivalent bundle or opportunity set. In the examples given in the previous subsection, they are evaluated by an equivalent lump-sum transfer $X. Instead of simply seeking to equalize $X perfectly, one can introduce the equivalent $X of every individual into a social welfare function, and proceed as if these values of $X were like utilities in the social welfare function.

These equivalent transfers $X can be used like utilities because they are faithful indicators of preference satisfaction. To see this, consider again the first example that involved zero work and the transfer $X. If the individual moves to another situation that he deems better, that will necessarily be equivalent to no work and a greater transfer. Therefore a better satisfaction is associated...
with a greater transfer. The same occurs with the second example that involves the possibility of working at the minimum wage. Again, this possibility has to be associated to a greater lump-sum transfer if the individual is more satisfied.

The fact that these are good indicators of satisfaction does not mean that they are just like the standard utilities of the welfarist approach. Recall that the standard utilities are unspecified. Their numerical calibration is a mystery in the welfarist approach, these functions have to be provided from outside the theory. In contrast, the equivalent transfers discussed here are quantities that can be computed on the sole basis of ordinal non-comparable individual preferences. Moreover, choosing one indicator (e.g., no work and $X) or another (e.g., working at the minimum wage and $2X) may be discussed in terms of the kind of equality one seeks to achieve, and how this favors individuals with particular preferences. A theory of the selection of the indicator is made possible with this approach, and it involves considerations of social equality and fairness, not an exploration of the nature of the human good. This feature of the approach echoes Rawls' claim that, for interpersonal comparisons, "the problem is not how to specify an accurate measure of some psychological or other attribute available only to science. Rather, it is a moral and practical problem." (1982, pp. 184-185)

Once such indicators of advantage are adopted, one can simply put them in a social welfare function and seek to maximize the value of social welfare by selecting the tax formula. Obviously, there remains the ethical issue of choosing the social welfare function. As it turns out, there seem to be good reasons in the theory of fair social orderings to give absolute priority to the worse off. These reasons have to do with satisfying simple transfer properties, such as the following: if two individuals have the same preferences, work the same amount, but have different levels of net income, it improves the situation to transfer some amount of income from the better off to the worse off. A social welfare function that measures individual advantage by an equivalent lump-sum transfer and does not give absolute priority to the worse off necessarily fails this property in some cases. This is because the transfer required by the property may be equivalent to taking a lot from the equivalent transfer of the better off and adding little to the equivalent transfer of the worse off. As the social welfare function measures their situation in terms of their equivalent transfers, this can be considered an improvement only if the worse off has much more priority. And as, considering all possible preferences in the population, there is no limit to the ratio between the loss of the better off and the gain of the worse off (in terms of their equivalent transfers), only an absolute priority for the worse off guarantees the desired property.

**Tax results**

We are now equipped with a well-defined social objective: maximizing the lowest individual value of the equivalent transfer, where the equivalent transfer is accompanied with zero labor (objective 1) or with the possibility to work at the minimum wage rate (objective 2). This objective differs from the Mirrlees approach only by the specific indicators of individual advantage, the specific "utilities" that are put into the maximin social welfare function. This is, in a sense, a small difference, but it makes it possible to deal with heterogeneous preferences without having to worry about arbitrary weights for different utility functions.

Maximizing this social objective when individual may widely differ in their preferences and in their earning abilities does not easily yield a general theoretical solution. This is due to the fact, which was explained in a previous section, that it is extremely hard to understand the structure of the set of incentive-compatible allocations in this context. Fortunately, a few things can be said nevertheless.

First, it is relatively easy to compare tax formulae, and therefore to address the problem of evaluating a reform. This problem is, in real political life, actually much more relevant than the quest for an optimal formula that, generally, stands no chance of being applied. There is even sometimes a danger in finding a general formula. When one finds, for instance, that the optimal formula involves decreasing marginal tax rates, it is tempting to conclude that any reform that moves in the direction of decreasing marginal tax rates would improve the situation. But, of course, nothing can be less true. The evaluation of a reform should not rely on a vague comparison with a stylized fact about the optimum. A rigorous and precise criterion must instead be used.

In the examples at hand, a precise criterion is easy to formulate when the population is sufficiently
diverse in preferences. For objective 1, a reform is a strict improvement if it increases the income support granted to those who have no earnings. For objective 2, a reform is a strict improvement if, over the range of earnings below the minimum wage, it reduces the greatest tax (or increases the lowest subsidy, when all levels of earnings in this range are submitted to negative taxes). These two criteria are valid under various conditions on the diversity of preferences. The simplest of such conditions consists in assuming that in the status quo (to which the reform is compared), there are some individuals with zero earnings, and some individuals who pay the greatest tax (or receive the lowest subsidy) in the low income bracket below the minimum wage.

Such criteria transform the problem of maximizing a particular social welfare function into that of maximizing (or minimizing) a particular point of the tax formula. This is very convenient for the evaluation of reforms, as it requires very little information about the characteristics of the population. The tax code itself suffices. Note that the objective of maximizing income support is congruent with the basic income idea promoted by van Parijs (1995).

One can also say something about the optimal tax. What can be said about the optimal tax for objective 1 is not very informative, but is better than nothing. It is that no individual should receive a greater income subsidy than those who have no earnings. This is quite intuitive, as giving such other individuals a greater subsidy than the basic grant would be a waste of resources, given that the worse off individuals, for objective 1, are those who have no earnings. This last point requires an explanation. For any tax formula, everyone is at least as well off (as viewed by her own preferences) as in the situation of earning nothing and living on the basic grant. This means that for everyone the equivalent transfer associated with no work is at least as great as the basic grant. And it is necessarily equal to the basic grant for those who are already in the situation of not working. This explains why the worse off, as measured by the equivalent transfer, are those who earn nothing.

For objective 2, a similar conclusion is obtained but it is more striking. The marginal tax rate can be set to zero for the low income bracket below the minimum wage. That is, all the individuals earning the minimum wage or less may receive the basic grant with full tax exemption for their earnings. Moreover, once again no one in the whole population should receive a subsidy greater than the basic grant. A tax exemption for low incomes is quite unusual in the Mirrlees setting with homogeneous preferences. This result is of course driven by the way in which individual indicators are computed, in relation to working at the minimum wage.

These results, clearly, do not say much about the optimal tax beyond the low income range. It is possible to do numerical simulations in absence of a theoretical formula for the optimal tax, with the computing power now available in computers. Such simulations have not yet been done for lack of data on the joint distribution of skills and preferences in the population. The traditional simulations in the Mirrlees tradition only require a distribution of skills and some estimate of the elasticity of labor supply, or some estimate of the average preferences in the population. Dealing more fully with the diversity of preferences is now possible in the fairness approach, but this is more demanding in terms of data.

Let us take stock. Compared to the Mirrlees tradition, the fairness approach to taxation can be viewed as providing a different kind of social objective, in which the exogenous utility function is replaced by an indicator of material advantage that reflects individual preferences but is based only on preferences and on nothing else about utility. The analysis of incentives is not changed at all, it only becomes more complex when the diversity of preferences is added to the inequalities in earning potential. The specific social objectives, which typically adopt the maximin form, deliver simple criteria for the comparison of arbitrary tax formulae, which may be very handy in the evaluation of reforms. The conclusions about the general shape of the optimal tax are so far limited, but nevertheless striking when they recommend a zero marginal tax rate for all incomes below the minimum wage.

**Limitations and extensions**

There remain some limitations to the approach. The theory of fair social orderings is often criticized for involving reference parameters, like the minimum wage in the case of objective 2. One could indeed take other values for the reference wage rate. But reference parameters are a virtue rather than a drawback of the approach. They
make it possible to develop an ethical theory of such parameters, and clarify the underpinnings, in terms of fairness, of the various measures of individual advantage. The limitation is, perhaps, that this ethical theory is somewhat less advanced than the rest of the theory. While, for instance, one understands rather well why an absolute priority to the worse off is obtained, the choice of a reference wage rate involves a series of considerations that are not homogeneous. For instance, it has been said earlier that taking a low wage rate as the reference protects the individuals with a high aversion to work. A related, but different consideration, is that taking a reference wage rate that is greater than the minimum wage rate will sometimes produce optimal allocations, in a perfectly planned economy with no incentive constraints, in which some work averse individuals are worse off than at the situation of zero work and zero consumption, a situation which is a violation of what is usually called the “participation constraint” (such individuals would rather opt out of the economy and migrate to a desert island).

Another consideration is that taking a wage rate that is below the minimum wage rate makes it impossible to satisfy the property that when all individuals have the same earning ability, the laisser-faire allocation is the best allocation for the social ordering. This property seems desirable if, in such a configuration in which individuals differ only in their preferences, all reasons to redistribute vanish. The literature also contains considerations formulated in terms of transfers from between individuals with unequal wage rates. In conclusion, the problem is not that a reference parameter has to be arbitrary, but that the theory of the good parameter appears to be work in progress.

More serious limitations have to do with the fact that the model is too abstract. This is also a criticism that can be leveled against the Mirrlees approach, of course. In particular, an essential feature of tax formulae, in practice, is that there is a different formula for different types of households. The model described deals only with individuals. Introducing households is a very complex issue when one deals with adults and children. The different needs of children must be taken into account, but there is no convenient basis for their evaluation in the theory of fairness. This theory is able to deal with any difference in needs that can be assessed by individual preferences. For instance, one can compare healthy individuals to paraplegic individuals if any individual can compare her situation to the other situation. Preferences may differ (and may be influenced by the situation one is in), but this is an issue that the theory is designed to deal with. But it seems much harder to ask an adult to compare his situation to that of a child, and to ask a child to compare his situation to that of an adult. There seem to be only two solutions to this difficulty. One consists in relying on the parents’ preferences over the situation of their children, thus considering children as a mere extension, or consumption, of the parents. The other consists in looking at individual situations from the lifetime perspective, but then the design of the tax becomes very complex.

Another feature of the existing theory that, it is usually said, calls for further explorations is that it assumes a stark difference between skills, for which individuals are not held responsible at all, and preferences, for which they are deemed fully responsible (although a concern for work averse individuals can be part of the theory). The suggestion is to explore variants in which individuals are partly responsible for their wage rate, or partly non-responsible for their preferences. I believe that one should be careful in developing such variants. If one follows a suggestion made in the last chapter of Fleurbaey (2008), there is nothing problematic in trying to respect preferences (when they are respectable), even if they are influenced by elements of the social background. So, the proper way to consider a partial responsibility for preferences might be either to be careful about the object of preferences or to correct preferences which are not respectable. The former case is relevant for instance for the assessment of the situation of individuals who are apparently work averse but in fact simply try to balance household duties with professional goals. Instead of saying that they are only partly responsible for their work aversion, one should correctly model the dimensions of their decisions, which involve not just consumption and leisure, but also the fulfillment of caretaking duties, for instance. Then one can seek to fully respect their preferences over this more complex depiction of their life.

The latter case (correcting preferences) occurs when one judges that the process of formation of preferences is too biased in favor of unfair conventions. For instance, if it appears that women are attracted by caretaking or homemaking activities because they are indoctrinated from youth.
that this is their social role, one can then seek to evaluate their situation with other preferences, even though it may be quite hard to estimate the proportion among them who are authentically attracted to become a caretaker and the proportion who are the victims of social manipulation.

But one should be careful about this sort of preference laundering. In the example of homemaking, one may wonder if women are pushed toward such social roles because they are less noble or if, conversely, they are considered less noble because they are traditionally the realm of women. There is probably no nobler activity than raising young human beings, for instance. So, perhaps in some cases the problem is not so much to correct preferences than to correct the disadvantages associated with certain objects of preferences. In the application of the taxation model, this is embodied in two options. Certain individuals who appear work averse may deserve to have their situation evaluated with more authentic, less work averse preferences – this the laundering option. The alternative is to consider that those individuals are actually working a lot, and are not work averse at all, but their work is not marketed and they are not decently paid! Then it is not difficult to reach the conclusion that they disadvantaged.

So much for preferences. As far as skills are concerned, it appears interesting to explore the fact that individuals have preferences not just about consumption and leisure, but also about their specialization, the contents of their job tasks. When individuals choose a specialization at the end of their studies, their consider earning prospects in different jobs and industries but also the genuine appeal of certain issues, certain forms of interactions with others, certain social roles. It would be very nice to enrich the model and to see what respecting individual preferences and choices over such choices would imply for income redistribution.

As explained in Fleurbaey (2008, chap. 5), if one assumes that individuals should not be held responsible for failing to go beyond high school, but can be held responsible for their wage rate beyond a certain level, then the worse off remain the unskilled in the analysis of taxation and nothing is changed to the evaluation of tax reforms and the design of the optimal tax, when the objective is the maximin applied to egalitarian-equivalent indicators of individual advantage.

Does the market have any moral value?

Luck egalitarianism and the market

Whether one refers to no-envy or to the egalitarian-equivalent approach, the theory of fair allocation revolves around the idea of equality of resources and gives a role to the market allocation. This role is somewhat muted in the case of egalitarian-equivalence, but not completely, and it is in particular possible to combine the egalitarian-equivalence approach to the measurement of individual advantage with the ethical requirement that, when all individuals are equally skilled and vary only in their preferences, the laissez-faire is the best allocation and there is no need for a tax at all (ignoring the costs of running the government and providing public goods).

This certainly goes against a narrow form of consequentialism that would evaluate social states only by looking at individual well-being, without any account of how individual situations are generated by transfers or by their own talent and effort. This feature of the economic theory of fair allocation seems well suited to philosophical developments in the same area. Indeed, all theories of justice that appeal to individual responsibility require going beyond narrow consequentialism and in the direction of laissez-faire. This is explicit in almost all of the main theories. Rawls is willing to rely on the market and on the income metric for the measurement of all-purpose means. Dworkin formulates the ideal allocation in terms of equilibrium of a particular insurance market with equal endowments. Arneson initially formulates the ideal of equal opportunity for welfare as involving a restraint on transfers: “Distributive justice does not recommend any intervention by society to correct inequalities that arise through the voluntary choice or fault of those who end up with less, so long as it is proper to hold the individuals responsible for the voluntary choice or faulty behavior that gives rise to the inequalities.” (1990, p. 176) Cohen similarly advocates that “we should ... compensate only for those welfare deficits which are not in some way traceable to the individual’s choices”. (1989, p. 914) Sen famously recommends looking at capabilities rather than achievements, which directly suggests looking at wealth rather than income, consumption, leisure. According to this school of thought, however diverse, a responsibility-sensitive form of
egalitarianism lies somewhere between a simple form of welfare egalitarianism and a simple form of libertarianism. The theory of fair allocation and its application to taxation seems well in line with this general approach. It seeks to respect preferences but at the same time has a very strong concern for equality in the means offered to individuals.

Murphy and Nagel do not say much about luck egalitarianism in general, but they discuss Dworkin’s approach, under the name of “equal libertarianism”, which seems apt (see p. 104sq). They correctly argue that this approach does not justify an equal sacrifice conception of fair taxation, but they seem to miss the point that it does involve giving some moral stature to market outcomes and transfers in the evaluation of the distribution.

The arguments in favor of market allocations

There are, however, two exceptions in the group of luck egalitarian theories that suggest an interesting debate could take place about whether transfers have any role to play in the evaluation of a social state. Arneson (2000) has proposed an alternative form of responsibility-sensitive prioritarianism that makes no direct recommendation about transfers, and simply gives greater priority to individuals who are worse off, who can benefit more from help, and who are less responsible for their disadvantage. Roemer (1998) has proposed a social welfare function that is strongly averse to inequalities due to circumstances for which individuals are not held responsible but behaves like sum-utilitarianism within subgroups of individuals sharing the same circumstances. These two approaches appear to locate luck egalitarianism somewhere between welfare egalitarianism and utilitarianism, rather than libertarianism.

Critics of market-oriented approaches generally argue that it is arbitrary to rely on market institutions because the only thing that matters is the distribution of well-being, not the process by which it is obtained. There is no reason, it is said, to give a special preference to market institutions as opposed to other forms of coordination of production and consumption. Such institutions are mere instruments in the service of well-being.

There is a grain of truth in this argument, but it involves a narrow form of consequentialism and confuses institutions with the outcome of institutions. At a deep level, no one (as least in the egalitarian forum) cares about the market as an institution. What some conceptions of justice care about is not the market itself, but the allocations that the market is able to generate. If some form of perfect planning was able to reproduce the market allocations without infringing on people’s autonomy, that would be fine. What is special about market allocations is not that they emerge from the market, but that they satisfy a certain pattern of activities and transfers, in relation to individual preferences.

The theory of fair allocation is helpful to understand how market allocations can be justified on the basis of extended forms of consequentialism. Consider first the no-envy property. It cannot be formulated with a classical social welfare function depending on individual utilities, because it involves the exercise of checking what utility everyone would get with the others’ bundles. This exercise has nothing to do with the market as an institution. But it turns out to be strongly connected to the idea of offering individuals the same set of bundles from which they can choose, and this, combined with some concern for efficiency, converges to a strong presumption that the market allocation (not the market itself) is particularly interesting.

Another notion that has played a role in the theory of fair allocation is the idea that when an allocation is obtained, nothing needs to be changed if preferences change in a way that everyone is even more satisfied than before with what he gets. This idea can be justified in terms of personal responsibility, because it means that nobody should be punished because he becomes even more satisfied with this current bundle, and it also means that the allocation that is selected should not depend too much on individual preferences. It can also be justified in terms of incentive compatibility, but this is not an ethical consideration and it will be left aside here. As it turns out, this weak form of independence of preferences, when it is combined with a concern for efficiency and a minimal form of impartiality, again delivers a strong presumption in favor of the egalitarian market allocation (not the market itself).

There are other justifications. For instance, take the fairness requirement that nobody should feel worse off than at the per capita endowment. This requirement is based on the idea that everyone has a kind of right to the per capita endowment,
and the allocation that is selected should be at least as good as that for everyone. This simple requirement, combined with a concern for efficiency and for subgroup consistency, once again vindicates the egalitarian market allocation.

The upshot is that, while it is true that the market institution is a mere instrument in the service of fairness and justice, it may happen that the market allocation, with its particular pattern of earnings and transfers, satisfies conditions of fairness that cannot be satisfied by a classical social welfare function that measures individual advantage in a narrow welfarist fashion. Fairness is more than a distribution of well-being narrowly defined.

Now, one may of course disagree with the notions of fairness that justify the pattern delivered by a market allocation. Having a preference for the egalitarian-equivalence approach over more market-leaning approaches in the context of income redistribution, because it is better for the neutralization of inequalities due to circumstances, I can only sympathize with a cautious attitude about market allocations. But one cannot claim that a moral preference for market allocations is completely arbitrary. It is debatable, but it has serious arguments. In particular, in the context of the Mirrlees problem of redistributing earnings, it is hard to resist the appeal of the laisser-faire when individuals have the same skills. This satisfies all the arguments listed here. Alternative approaches generate envy, make some people worse off than at the average bundle, and are excessively sensitive to small changes in people’s preferences. None of these considerations is fully compelling, but they add up to a strong presumption in favor of the (perfectly equal) laisser-faire.

**In the second best, all approaches are market-oriented**

There is a French proverb that says that all cats are gray in the dark. It is an ironic twist to the ethical debate about the moral stature of the market allocation that incentive constraints tend to blur the difference between market-oriented criteria and other criteria.

Take the result that the marginal tax rate should be null in the low income bracket (below the minimum wage). This property of the tax formula, which in a sense means that market earnings will be respected (up to a lump-sum transfer, because everyone gets the basic grant), is obtained with a certain specific egalitarian-equivalent criterion, one that takes the minimum wage rate as the reference. It is also obtained with a wider class of criteria that more heavily lean on the market and involve a general preference for submitting all individuals sharing a same skill level to an equal amount of transfer. The literature on the fair compensation of unequal skills has highlighted the difference between the egalitarian-equivalence approach, which is good at neutralizing inequalities due to skill differences, and the alternative approach that focuses on the concern about equal transfers for equally skilled individuals. But this difference, which appears clear-cut in axiomatic analysis, becomes much less stark when one looks at the best taxes, that is, when one takes incentive constraints into account.

Another example of the graying of ethical differences under incentive constraints is offered by Roemer’s examples of optimal taxation. Roemer often takes specific forms of utility functions that are called quasi-linear in economics (see, e.g., Roemer 1996, pp. 297-301). Such utility functions represent preferences that have the special property that the willingness-to-pay for extra leisure does not depend on one’s level of consumption. Moreover, with such utility functions, utility is equal to net income, and can be measured in dollars, with a deduction for the “cost” of working. This measurement of utility, for such special preferences, is exactly identical to the transfer equivalent with zero work that was defined in a previous section (the difference is that the transfer equivalent can be computed for all kinds of preferences).

It turns out that when all individuals have the same skills, which is a case in which Roemer’s approach does not recommend the laisser-faire but the maximization of the sum of utilities, the best allocation for the utilitarian sum is the laisser-faire. So, this is another example in which the difference between the market-oriented approaches and a non market-oriented approach vanishes.

This does not mean that ethical differences no longer matter. There are cases such that,
even under incentive constraints, the various approaches described here deliver different policy recommendations. It is not the same to advocate maximizing the basic grant (recommendation from objective 1), the uniform grant given to low incomes (recommendation from objective 2), the average utility of the unskilled (recommendation from Roemer’s approach), or the more complex formula that emanates from the more market-oriented luck egalitarian criteria. There is still some room and relevance for moral debate.

A brief conclusion can be offered here. This was meant to be a defense of the economists’ work on taxation. Economists are less naïve, and less libertarian, than suggested by Murphy and Nagel. They are divided into welfarists (Mirrlees’ tradition) and luck egalitarians (the fair allocation tradition). The former are probably closer to what Murphy and Nagel would like to see, but the latter are closer to what political philosophy has been advocating over the last two decades. Perhaps it should be stressed, finally, that Murphy and Nagel’s critique of conservatism in taxation debates is excellent, and that their book was taken as the sparring partner of this chapter just to say that economists are actually on their side of the barricade.

References
**Dernières parutions**

**Working Papers**


Guilhem Fabre & Stéphane Grumbach, *The World upside down, China’s R&D and innovation strategy*, FMSH-WP-2012-07, avril 2012.


**Position Papers**


Informations et soumission des textes : wpfmsh@msh-paris.fr

http://halshs.archives-ouvertes.fr/FMSH-WP

http://wpfmsh.hypotheses.org